Draft

Drift of Emissions Measurements in the EPAct Program

April 30, 2008
George Hoffman

Drift might impact EPAct significance testing of emissions data by increasing variability and/or by biasing results, changing the likelihood of obtaining statistically significant results. Before considering how to deal with drift in EPAct, we first must consider standard engineering practice, the evidence that drift exists, what can be done about drift, and the actual definition of drift.

Standard engineering practice is an important guide in determining how to test and control for drift during the course of emissions tests. Drift, discussed below, is a change in emissions values after repeated measurements or after a period of time.

The following thoughts are based on no data, although data from an existing study (E-67) might be used to test for drift as a function of time. How E-67 data might be used to study the existence of drift is discussed below.

Data have not been presented on such issues as the mechanism of how drift is produced in a study and the impact of drift. Hence, the following treatment at this time is more speculative than quantitative.

What is drift? At least two kinds of drift are possible.

Random drift—mean emissions do not change between the first and second emission measurements, but the variance of the second replicate around the mean is larger than would be expected by chance. This would increase the error of measurements and possibly result in greater difficulty in obtaining statistically significant results.

Biasing drift—mean emissions change between the first and second measurement. The two measurements for a vehicle-fuel would not be true replicates because they do not share the same mean and they might also have a different variance.

Obviously there is random error in any emissions measurement. Changes in variables such as temperature or humidity can impact measurements, but these impacts are presumably random and do not result in a change of mean and standard deviation from replicate to replicate, regardless of when the replicates occur.

Notwithstanding the above discussion of drift, is there an existing, established definition of drift? The above, proposed definitions are presented for the purposes of this paper.

Follow standard practices whenever possible.

Any information on standard practices should be used to determine the appropriate approach to drift for EPAct data collection. Data from E-67, or other sources if available, should be obtained to investigate the reality of drift for Tier 2 emissions studies.

Possible use of the study, E-67, as a source of drift information.

Durban et al., 2006, Effects of Ethanol and Volatility Parameters on Exhaust Emissions, CRC Project No. E-67, provided at least two replicates per vehicle-fuel. The order of fuels tests was randomized for both the first and second replicate, which was performed only after the first-replicate tests were completed for all fuels. Also, 43 third replicate tests of 331 total tests were performed.

If drift depends on the time between tests, then a test for drift might be devised based on replicate tests on the same vehicle-fuel, closer together in time, compared with replicates further apart in time. This assumes that drift is proportional to the time between successive tests. Differences in variability might also be considered.

EPAct studies will have a different procedure than E-67. Each vehicle-fuel combination will be tested twice in succession, and a third test performed if the two tests produce sufficiently different emissions. Drift here would relate to biasing or variance changes in emissions measurements for a vehicle-fuel. Possibly drift could be considered for each vehicle-fuel but more likely drift would refer to emissions changes that might occur at least partially based on the time frame of the experiment. Do the measurements of emissions for a vehicle-fuel test depend on whether the tests were performed early in the experiment or later in the experiment? Possibly the E-67 study could provide an answer to this question, at least for the protocols used in the E-67 study. Of course, the interpretation of these results is hampered by the lack of a mechanism for drift.

The E-67 paper asserts “the data showed very little confounding of test sequence with the variables under study and indicated that the randomization has no substantial problems.” From this, it is unclear that Durban et al. tested specifically for drift, but such a test might well have been carried out. Certainly the correlation of the time between successive tests and changes in emissions could be tested. Depending on the data, E-67 results might be applied to the specific question on drift asked for the EPAct study: do average emissions for a vehicle-fuel depend on when in the study the emissions are measured?

Drift, even when it is found to exist, might be ignored.

Drift would be very difficult to detect in studies with successive vehicle-fuel tests. Procedures similar to those used in the E-67 study might be applied to a study of drift but such replicate-test procedures are expensive to implement.

If drift is detected, it is unlikely that anything can be done about it. When drift occurs, what approaches can be applied to correct for it? Some techniques might be applied, assuming drift could be corrected by a simple mathematical relationship. Specifying and testing empirical corrections would probably require a large amount of data.

A mechanism to explain biasing drift would be helpful. Possibly a mechanism could allow a correction for drift while data are collected. If data are collected with an understanding of what causes the drift, a subsequent emissions models and testing might

take the effect of drift on emissions into consideration. Without a mechanism of action and a clear understanding of drift, direct data corrections for a drift in emission values over time should not be attempted.

If drift is shown to occur, does this imply that emissions measurements are erroneous? Depending on the mechanism(s) of drift, a more dynamic model of emissions might be necessary, to account for the changes in vehicle-fuel emissions over time.

Drift might be difficult and expensive to detect. If the impact of drift potentially biases emissions studies, possibly it could be detected in studies, such as E-67 employing Tier 2 data. Another possibility is that drift would be difficult to detect, requiring a considerable data collection effort. EPAct sample size requirements were based on detecting a 25% difference in emissions from fuels in two replicates for 19 vehicles. To detect drift, the two replicates would have to be collected at different times and not be collected successively, one after another. If drift produces a bias of 25%, it probably would have been raised as a serious issue before now. If drift produces relatively small differences in emissions (much less than 25%), it is very possible that statistical techniques will not reliably produce significant differences between replicates subject to drift, unless massive numbers of replicates are taken.

How often is data collected on drift in emission studies? Has anyone studied the cause(s) and impact of drift on emissions? How large is the drift effect? If this information is available, it would be valuable input to the current discussion.

Summary and some recommendations.

Engineering practice generally should be followed if a consensus exists on approaches to drift, even if that consensus is to ignore drift. If there is no current, consistent, applicable engineering practice, a decision has to be made on how to handle drift. If there is a consensus on an engineering practice concerning drift, the EPAct approach should follow the engineering practice unless there is a very good reason not to follow the consensus.

E-67 and other studies, if available, should be used to investigate drift. This study might indicate whether significant biasing or changes in variance occur with time. If there are no detected significant effects of drift, then there is no evidence that drift occurs in a carefully conducted experiment. To look for subtle effects of drift on emissions would be costly and possibly not worth the effort. What is the evidence that drift produces significant and meaningful changes in emissions over the course of an emissions experiment?

So long as there is no consistent engineering practice on drift and if appropriate Tier 2 studies (E-67) produce no evidence for meaningful changes in drift, no effort should be expended to study drift. Why expend resources on a complicating factor that has not been proven to be a factor in emissions testing? If, however, there is proof that meaningful drift exists, then this analysis must be modified to include this proof.

Of course meaningful drift, if it exists, might affect the statistical tests of studies, especially Tier 2 studies such as EPAct. Inability to produce statistical significant results in a large study would seriously compromise the utility of the work. So if there is reason to believe that "biasing drift" occurs in a carefully designed experiment, mitigating the effect of drift would be worth considering. "Random drift" might be possible to include in the RANDOM statement of a mixed model.

If meaningful drift is found or assumed, then the question naturally arises, what to do about drift? If drift is found in E-67 data, it might require considerable effort, and a change of methodology, to test for drift in EPAct. If found, a model for drift would have to be produced and possibly applied to the data analysis (something like Analysis of Covariance or ANCOVA) and this might not be feasible.

One other, obvious approach to drift is to ignore it. Assume only that the emissions results correctly and accurately measure emissions. Assume further that drift produces only a small contribution to random error. "Biasing drift" is a possibility but it is not proven for emissions studies, as far as I know. (Please provide proof for biasing drift if such proof exists.)

No data on drift has been presented to prove the drift can be important to emissions measurement and to produce a reliable estimate of the impact of drift on emissions. No factors or agents have been identified that produce drift. No quantitative information has been presented on the size and occurrence of drift. Without knowledge of what causes drift, would any results or formula be applicable to drift in future emissions studies?

In summary, drift is potentially important to Tier 2 emissions studies, including EPAct. E-67 data should be tested for evidence of drift. If no evidence is found in E-67 or other data (especially Tier 2 data) and if standard engineering practice does not include consideration of drift, then drift should be ignored. Otherwise, an attempt should be made to include terms that allow for drift in the actual data analysis. If data indicate that drift remains a significant problem that can not be solved in the data analysis, a separate study of drift might be required. My opinion is that such a study of drift would require considerable resources. The actual sample size computations for a study of drift would require information on the means and standard deviations of drift.